

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/349073738>

# An investigation into the WIV databases that were taken offline

Preprint · February 2021

DOI: 10.13140/RG.2.2.28029.08160

CITATIONS

3

READS

53,481

3 authors:



Charles Small

5 PUBLICATIONS 3 CITATIONS

SEE PROFILE



Billy Bostickson

DRASTIC

21 PUBLICATIONS 8 CITATIONS

SEE PROFILE



Gilles Demaneuf

Engineer and Data Scientist

30 PUBLICATIONS 12 CITATIONS

SEE PROFILE

## DBs Taken Offline

### Objectives:

- List all the DBs that were taken offline
- Explain their significance
- Back it up with screenshots and proofs of deletion

## 1. WIV: Bat and rodent borne viral pathogens - batvirus.whiov.ac.cn

### 1.a Findings:

- Batvirus.whiov.ac.cn is an essential database
- Batvirus.whiov.ac.cn fits into key Chinese and international programs
- Part of batvirus.whiov.ac.cn data was always confidential
- Partial funding of samples collection and sequencing via the NIH does not secure ownership of resulting data
- The aim of the projects funded by the NIH includes pathogenicity enhancement studies using unpublished viruses
- Financial support and membership of EVAg is at odds with the unavailability of batvirus.whiov.ac.cn and the lack of transparency of the WIV
- External access to Batvirus.whiov.ac.cn ended on 12th Sep 2019
- Untimely changes to the database description on 30<sup>th</sup> Dec 2019 may betray some sensitive research
- Batvirus.whiov.ac.cn holds essential information on SARS-CoV-2 origins
- Questions as to the unavailability of batvirus.whiov.ac.cn are not being properly answered
- The suspension of the availability of batvirus.whiov.ac.cn is contrary to Chinese law

---

### Finding: Batvirus.whiov.ac.cn is an essential database

- The administrator of the database is Shi Zhengli herself.
  - In <http://batvirus.whiov.ac.cn> (the WIV address for the database), whiov stands for Wu Han Institute Of Virology. The 'batvirus' prefix can be confusing as this is not just a bat database despite the name.
  - The alternative address is <http://www.scidb.cn/dadaSet/handle/768> which is its online location on a Chinese service for science datasets.
- It is the most important BatCoV database in China and holds the samples and sequences from the WIV. A description is available [online here](#) (English [translation](#)).
- It focuses primarily on bats and rodents, but also held bird and arthropods (insects & ticks) samples
  - The database held more than 22,000 samples and sequence records, including for all WIV sampling trips going back many years, plus some data from external DBs (namely [DBatVir](#) (bats) and [DRodVir](#) (rodents), managed by the Institute of Pathogen Biology in Beijing - both still available).
  - Tick pathogens were also very much a field of study, with more than 50,000 tick samples collected.
- The database contains data on seasonal epidemics of viruses crossing the species barrier, data not provided by DBatVir (for bats) or DRodVir (for rodents). This makes it potentially the best database for investigating whether the theory of natural spillover of SARS-CoV-2 is plausible. The database includes samples and sequences of bat beta-coronaviruses gathered by the WIV on trips to Yunnan including Mojiang (where RaTG13 was sampled), sequences that have not all been made public.

- Rodents were a common WIV sampling target too. One of the first sampling trips by the WIV at the Mojiang mine in 2012 (shortly after the death of 3 workers who removed bat guano there) would include sampling of rats, not just bats. On that occasion the [Mojiang paramyxovirus](#) was discovered.
- More specifically:
  - The WIV had collected over [15,000 samples from bats](#) (amongst the 22,000 records of samples in the DB)
  - The DB contained over [1,400 strains](#) of viruses across all species (animal, insect), including around 1,000 coronaviruses, with at least [500](#) recently discovered bat coronaviruses, and at least [50](#) of these close to SARS.
  - At least 19 of these coronaviruses were isolated (meaning grown in cell culture).

#### **Finding: [Batvirus.whiov.ac.cn](#) fits into key Chinese and international programs**

- The sampling, sequencing effort and building of the database was funded under program [2013FY113500 "Investigation of viral pathogens of major natural hosts and vector insects in China"](#), launched by the Chinese Ministry of Science & Technology in May 2013, a program that was being reviewed around the time of the outbreak.
- The database also aligned perfectly with the successor program "[Scientific survey of principal natural viral pathogen resources in China](#)" that was put out to tender in July 2019 by the Chinese Ministry of Science & Technology (thereafter the "**2019 Scientific Survey**"). That successor program [was to be led by the Shanghai lab of Professor Zhang Yongzhen](#), who was the first to publicly share a SARS-CoV-2 sequence. Employees of the Shanghai lab confirmed that SARS-CoV-2 was identified under this successor program ([1](#), [2](#), [3](#))
- **Note:** There were later unverifiable stories that the program (likely meaning funding under the program) was delayed pending final approval. These vague stories could very well be disinformation to try to dissociate any recent research from the very specific research objectives of the 2019 Survey. In any case the 2019 NIH-EcoHealth grant already provided some funding.
- The new program is in line with the [China National Global Virome Initiative](#), part of the China Virome Project (launched Feb 2018 at the Prince Mahidol Annual Conference in Bangkok), which itself is the Chinese implementation of the international [Global Virome Project](#) (launched Aug 2013)
- The [proof-of-concept](#) of the GVP was the Emerging Pandemic Threats programs managed by USAID/PREDICT (EPT-1 started 2009 for 5 years, then [EPT-2](#) started in 2014).
  - EPT-2 was [scaled down in Jan 2019](#) by [the Trump administration](#), 8 months before it [reached the end of its 5-year funding cycle](#) (Sep 2019). It nevertheless received an administrative extension (with no budget) to March 2020 to be able to write up final reports.
  - The scaling down of EPT-2 resulted in an intense politicization of the debate, still evident to this day.
  - [batvirus.whiov.ac.cn](#) was in line with the [proposed research direction](#) recommended by WIV researchers in March 2019 to investigate cross-species transmission and human pathogenesis of bat SARS-related coronaviruses, with the inclusion of rodents for instance.

#### **Finding: [Part of batvirus.whiov.ac.cn](#) data was always confidential**

- [Batvirus.whiov.ac.cn](#) holds a password protected section for data that was not to be freely shared. That private data includes:
  - (1) Sampling data of viruses not yet sequenced and (2) sequences of viruses that have not yet been the object of publication.
    - **Note:** The idea is very likely to get the WIV and its associates the privilege of writing the first papers on the viruses sampled by the WIV.
  - (2) sampling data from sites that the WIV prefers not to make public.

- **Note:** This may be to preserve the WIV primacy and avoid having teams from other labs studying BatCoVs sampling there and potentially writing papers ahead of the WIV on viruses that may be found in these sites.
- As [per the description](#) (English [translation](#)):

Due to the use rights of some data (referring to the cases where the virus sequence has not been published, the virus sequence has not been uploaded to NCBI, [and the wild animal sample information of the sample collection site cannot be published](#)), if the user needs to access and use this part of the data, you can by contacting the relevant management personnel of this database, obtaining the login account password of the platform through identity verification and authentication, and logging in to the platform to use relevant data.

### **Finding: Partial funding of samples collection and sequencing via the NIH does not secure ownership of resulting data**

- Key aspects of the WIV research were partially funded by the NIH/NIAID via EcoHealth.
- A total of \$3.6 mln was thus granted to EcoHealth [through a succession of yearly payments](#) under the NIH project 'Understanding The Risk of Bat Coronavirus Emergence'.
- Part of these grants were then spent on research projects with the WIV, [for a total of \\$598,500](#) paid to the WIV as fees for the [stated purposes](#) of:
  - conduct[ing] high-quality testing, sequencing, and analyses of field samples;
  - maintenance of cold-chains from field to lab;
  - ensuring quality control of sample storage and testing;
  - collaborating on scientific publications and programmatic reporting.
- These purposes thus include the collection and analysis (sequencing) of viral samples, whose results are stored into batvirus.whiov.ac.cn
- Interestingly these purposes also include quality control and safety issues.
- These US-funded samples and sequences were not automatically sharable - [China could typically decide what it wanted to share](#). USAID had effectively no clear ownership of the data, despite funding the EPT programs
- This was clearly highlighted [as an ambiguity and a potential issue](#) in a USAID evaluation of EPT-2:  
[underlining added]

**“USAID does not own data collected under PREDICT 2:** Information from interview respondents familiar with PREDICT 2 (including ETD respondents) and a review of the procurement documents provided to the evaluation team suggest that PREDICT 2, not USAID, holds the raw data it collects for its research and models. [This dramatically constrains the control over those data for future programing, research, and dissemination](#). An ETD respondent explained that USAID would have access to the data, and noted that “PREDICT plans to upload [the findings approved for release \(only a subset of samples in four countries so far\)](#) to its global website in the near future” and that they will be available to USAID.

[---]

In identifying viral pathogens, PREDICT 2 has strict protocols for protecting country data. Approval from three pertinent ministries (usually health, livestock, and wildlife) is necessary before [findings](#) are released. The evaluation team commends this approach as an avenue for ensuring country ownership. In terms of [raw data](#) collected under the project, PREDICT 2 says “the PREDICT Consortium holds the raw data collected, though USAID has access to raw data through the Agreement Officer’s Representative [AOR].” While the evaluation team is not qualified to provide a legal opinion on data rights and ownership, we believe clarity is required with respect to USAID’s legal authority over data collected, as opposed to “accessed” under PREDICT 2, [given the finite nature of the project used to collect the data and the possible future uses for those data](#)”

- In other words there is a legal ambiguity by which at the end of the EPT-2 program, USAID (and hence the US taxpayers) may not have legal access to the data from the sampling they partially funded (on top of the country's right to oppose the release of findings).
- Practically this means that
  - there is no legal way to force China to at least release the data that was collected with the financial help of USAID-PREDICT
  - with the passing of EPT-2 the ownership of non-disclosed EPT-1 and EPT-2 data may now actually fully reside with China.
  - Peter Daszak, president of EcoHealth Alliance which was involved in the EPT samplings with the WIV, said just as much in an interview:

“The NIH has told us not to work on this project. Obviously, we're not going to break any NIH rules. But we have an ongoing collaboration, we have data that we've gathered over 15 years of working in China — 5 years under a previous grant from the NIH — which haven't been published yet. So we need to carry on with that work.”
- **Note:** This is particularly important for the WIV/PREDICT sampling trips around Yunnan such as the 2012 sampling trip at the time of the 'miners' disease in the abandoned Mojiang mine and many of the subsequent sampling trips (some of which are available on data.predict.global).

### **Finding: The aim of the projects funded by the NIH includes pathogenicity enhancement studies using unpublished viruses**

---

- The actual virus research mentioned in the NIH grant was conducted in the WIV labs in cooperation with WIV researchers, and such cooperation meant alignment of research objectives between the NIH grants and the WIV.
- We can indeed see that research objectives of the EcoHealth projects detailed in the NIH grants are in line with the objectives of the Chinese programs 2013FY113500 and its successor, the '2019 Scientific Survey'.
- The objectives of the Chinese '2019 Scientific Survey' have a focus on studying biosafety risk (pathogenicity) as the cell level (cell cultures / in vitro) and at the small animal level (humanized mice / in vivo), using new viruses - meaning unpublished viruses kept in the private section of batvirus.whiov.ac.cn:

[15] (4) Complete the pathogenic characteristics analysis of 10 important new viruses/strains, which for at least 5 important pathogenic new viruses should be based on a biosafety risk assessment at the cell level and small animal level.

- This, as expected, is in line with the descriptions of the virus experiment projects in the EcoHealth NIH grants:
  - The latest NIH grant was for project 2R01AI110964-06, with a budget start date of 24<sup>th</sup> July 2019 and a budget end date of 30<sup>th</sup> June 2021. The objectives of that latest project include:

Aim 3. In vitro and in vivo characterization of SARSr-CoV spillover risk, coupled with spatial and phylogenetic analyses to identify the regions and viruses of public health concern. We will use S protein sequence data, infectious clone technology, in vitro and in vivo infection experiments and analysis of receptor binding to test the hypothesis that % divergence thresholds in S protein sequences predict spillover potential.
  - The previous grant for project 5R01AI110964-05 (start date: 1<sup>st</sup> Jun 2018, end date: 31<sup>st</sup> May 2019) which included in its objectives:

3. Test predictions of CoV inter-species transmission. Predictive models of host range (i.e. emergence potential) will be tested experimentally using reverse genetics, pseudovirus and receptor binding assays, and virus infection experiments across a range of cell cultures from different species and humanized mice.

**Finding: Financial support and membership of EVAg is at odds with the unavailability of batvirus.whiov.ac.cn and the lack of transparency of the WIV**

---

- The WIV is a [participant in EVAg](#) (European Virus Archive (goes) Global) - a [5-year international cooperation project](#) funded under the European research and innovation programme 'Horizon 2020' which has been succeeded by 'EVA Global'.
- Under that cooperation project the WIV has already received and may still receive European money in the form of two grants (under EVAg and EVA Global). The [total amount deliverable](#) to the WIV under the 2 grants is around Euro 220k.
- The [stated objective](#) of EVAg is to '*generate the largest collection of mammalian viruses in the world and move beyond the current state-of-the-art to provide an increasingly valuable resource and service to the world's scientific community*'. Hence the non-availability of the fundamental WIV database is fully at odds with the objectives of EVAg, to which the WIV is a participant and as such a recipient of European grants.
- This has been duly noted by the European Commission; on the 9<sup>th</sup> Dec 2020 [in response to a question by a Dutch member](#) of the European parliament, [Commissioner Gabriel reported that](#):

"a lack of communication by WIV was observed in the period from February to October 2020" and that "for this reason, an interim report on its activities has been requested".

**Finding: External access to batvirus.whiov.ac.cn ended on 12<sup>th</sup> Sep 2019**

---

- Batvirus.whiov.ac.cn had been online for a few years, saw a version 2 released in June 2019, went inactive for a week [during the second half of August 19](#), before [becoming definitely inaccessible](#) (out of the WIV at least) on the 12<sup>th</sup> Sep 19. It was online intermittently after this date [from mid-December 2019](#), and occasionally [until February 2020](#), but was not accessed from outside of the WIV after 12 September 2019.
- While the non-WIV DBatVir and DRodVir components can still be accessed independently, all the WIV generated data, and especially the unpublished data in the password protected section was thus rendered inaccessible.
- No reason can be found for the week of inactivity in August.
- There have been repeated contradictions by key people about the reasons for the final removal of the DB in Sep 2019, and about the exact content of that DB:
  - In Dec 2020, Pr. Shi Zhengli explained [in a BBC interview](#) that access to the DB was [stopped](#) to prevent cyber security attacks.
  - On the 26<sup>th</sup> January 2021, Pr. Shi Zhengli confirmed again that the database has been taken offline 'during Covid-19 pandemic' in an email answer to [Tommy Cleary](#).
  - On the 10<sup>th</sup> March 2021, during a [Chatham House interview](#), Peter Daszak repeated the exact reason given by Shi Zhengli in her email to Tommy Cleary above.
    - **These three statements do not make any sense since the main database was taken offline on the 12 Sep 2019, 3 months at least before the official start of the pandemic.** So either the reason given for taking the database off is not correct (which raises more questions), or the statement points at an outbreak in Sep 2019.
    - In any case, the 61.5MB MySQL DB could easily be hosted by safe servers in China, segregated from the WIV network. Databases of similar size are widely available [on the CSData website where the description is located](#). It should have been possible in any case to simply hand over a USB key or drive with the DB.
  - In the same email to [Tommy Cleary](#), Shi Zhengli asserts that all sampling details sequences contained in the database (partial or full) have been already published
    - **This is in full contradiction** with the description of the DB and the existence of a password protected section for unpublished viruses. [See description](#) (English [translation](#)):

[Due to the use rights of some data \(referring to the cases where the virus sequence has not been published, the virus sequence has not been uploaded to NCBI, and the wild animal sample information of the sample collection site cannot be published\), if the user](#)



needs to access and use this part of the data, you can by contacting the relevant management personnel of this database, obtaining the login account password of the platform through identity verification and authentication, and logging in to the platform to use relevant data.

- This is also in contradiction with a [message from P. Daszak](#) on the 10<sup>th</sup> Dec 2020 where he wrote that:

"Thanks to @RudyGuliani's poorly thought out interference, a grant to @EcoHealthNYC funded under @realDonaldTrump was terminated by @NIHDirector & now we can't get access to critical samples that would help us understand the origins of COVID & could be used to improve vaccines'

and also with a previous [interview with NPR](#) in April 2020, where P. Daszak explained that:

"You can't just turn up as an American and say, 'I want to find out what viruses you've got,'" says Daszak. "You have to work with local collaboration and with the permission of the governments."

- Further, during the [Chatham House interview](#) of the 10<sup>th</sup> March 2021, Peter Daszak explained that the international part of the joint-team did not request to have access to that main database, because he told them that he essentially knew what was in that DB and that there were no virus closer to RaTG13 there.
  - This does not align with previous statements where Peter Daszak reiterated that without cooperation with the WIV, foreign scientists could not get access to the private section of the DB (especially for the extensive sampling work done directly by the WIV or CDC, without Echo Health Alliance/PREDICT involvement on the ground). See statement above for instance.
  - This also ignores the key fact that the very purpose of the current research at the WIV was to study pathogenicity of new (unpublished) viruses, as detailed in the [next section](#).

Per [2019 Scientific Survey](#) (initiated July 2019):

[15] (4) Complete the pathogenic characteristics analysis of 10 important new viruses/strains, which for at least 5 important pathogenic new viruses should be based on a biosafety risk assessment at the cell level and small animal level.
  - Incidentally, in the same Chatham interview, Peter Daszak incorrectly described the DB as being an Excel spreadsheet when it is actually a relational database (a [61.5MB MySQL DB](#)).

### Finding: Untimely changes to the database description on 30<sup>th</sup> Dec 2019 may betray some sensitive research

---

- While the database itself was not accessible outside of the WIV after the 12<sup>th</sup> Sep 2019, its description factsheet still was accessible until fairly recently.
- For some unexplained reason, on the 30<sup>th</sup> Dec 2019 - the exact day whistleblower doctors in Wuhan raised the alarm - a new version of the database factsheet was uploaded at the time Pr. Shi Zhengli was on the Shanghai to Wuhan high speed night train. The changes are [highlighted here](#).
  - We do not know whether the DB data itself was updated on that occasion - the data it contained had not been accessible for external parties since the 12<sup>th</sup> Sep 2019 in any case.
  - The writer of the factsheet may be the WIV researcher Li Bei, as she is described as being responsible for the text writing in the factsheet itself. She may have uploaded the new factsheet herself or Pr. Shi Zhengli may have done so while on the train.
- It is difficult to understand why Pr Shi Zhengli would push an update of just the database description on such an eventful day, if there wasn't a good reason:

'The mysterious patient samples arrived at the Wuhan Institute of Virology at 7 P.M. on December 30, 2019. Moments later Shi Zhengli's cell phone rang. It was her boss, the institute's director. The Wuhan Center for Disease Control and Prevention had detected a novel coronavirus in two hospital patients with atypical pneumonia, and it wanted Shi's renowned laboratory to investigate. If the finding was confirmed, the new pathogen could pose a serious public health threat—because it belonged to the same family of viruses as the one that caused severe acute respiratory syndrome (SARS), a disease that plagued 8,100 people and killed nearly 800 of them between 2002 and 2003. **"Drop whatever you are doing and deal with it now,"** she recalls the director saying.'

Source: [Scientific American](#) - originally published 24<sup>th</sup> Apr 20.

- One possible scenario is that she updated the DB itself and at the same time pushed an update to the description that was anyway pending.
- Another scenario is that the timing of the changes in the description is intentional and meaningful. As discussed below (Section 1.b). In that line one could imagine that the changes may try to dissociate the database from its relation to the [2019 Scientific Survey](#) (initiated July 2019) as the objectives of that project include an **item that should raise attention**:

[15] (4) Complete the pathogenic characteristics analysis of 10 important new viruses/strains, which for at least 5 important pathogenic new viruses should be based on a biosafety risk assessment at the cell level and small animal level.

- That is indeed a lot of dedicated pathogenicity work on likely undisclosed viruses (that's very likely what new means - publication if any would follow the pathogenicity work)
- Good candidates for new viruses/strains could well include some of the recently mentioned 8 BatCoVs from the same clade as RaTG13.
- Additionally it is advanced work using humanized small animals (likely mice).
- The Jul 19 date for the inception of the tendering process for the 2019 Scientific Survey in China does not mean that no such research took place before that at the WIV or elsewhere. This kind of pathogenicity research was not particularly new to the WIV. In fact the 2019 Scientific Survey just builds on and focalizes existing work and provides additional funding for more of that work - right at a time when the US were cutting their own funding through USAID-PREDICT EPT.

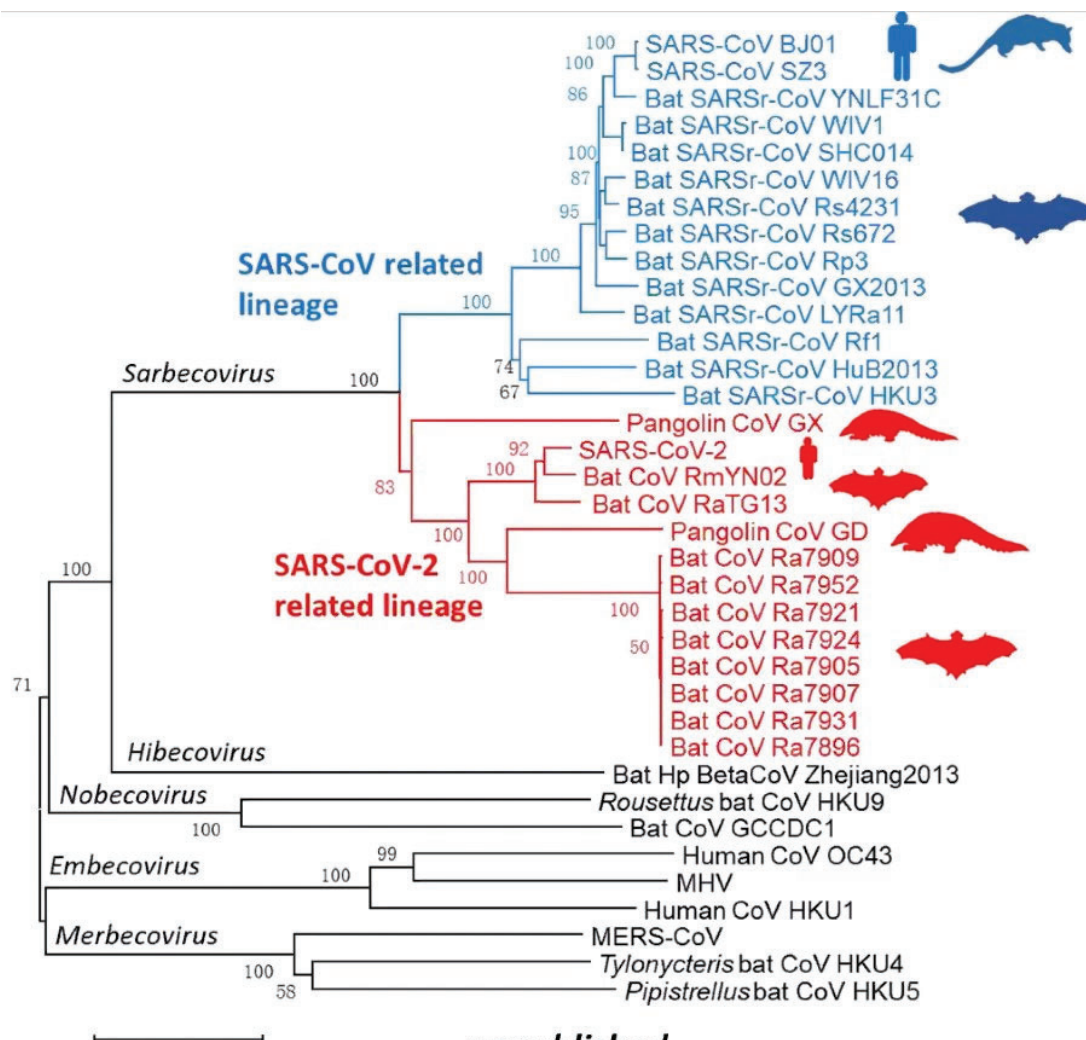
#### **Finding: [Batvirus.whiov.ac.cn](#) holds essential information on SARS-CoV-2 origins**

- There are estimated to be at least 100 unpublished sequences of bat betacoronaviruses in [Batvirus.whiov.ac.cn](#), which urgently need to be accessed by international scientists in order to investigate the origins of SARS-CoV-2:
- These include in particular:
  - WIV6 and WIV15 BatCov isolates which have never been made public, but whose existence can be deduced from gaps in the existing series of isolates mentioned in research papers.
  - 8 beta BatCoVs related to SARS, sampled from Yunnan, only recently mentioned by Shi Zhengli in an [Addendum to a Nature paper](#) and in an [online presentation](#) and never published despite being of the highest relevance.
    - They are likely the same as the 8 from the 7896 clade published [as part of a batch of 630 viruses of Latinne et al](#) without further details. On that occasion only the RdRp portions were uploaded to GenBank on the [13 Aug 2019](#) (around 360 bases, not the full sequences of the viruses).
    - That clade is very close to RaTG13 (the closest relative of SARS-CoV-2), so should offer some very important clues as to but is not discussed in the literature.
    - Questions as to why no more attention is being paid to that essential clade [are not being answered](#).
  - These 8 plus WIV5 and WIV6 may particularly make good candidates for the objective 15.4 of the 2019 Scientific Survey listed above.



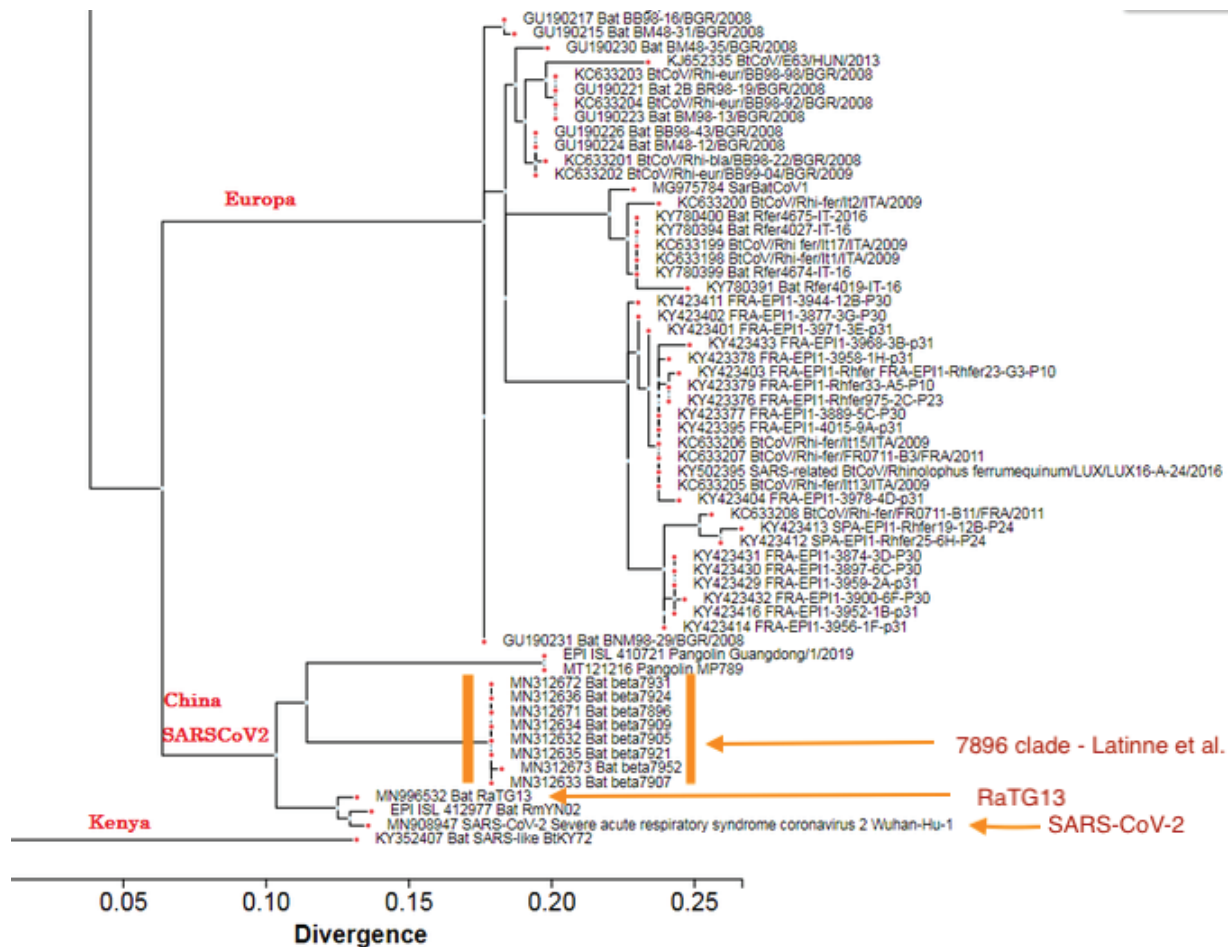
- This also includes all the undisclosed BatCoV samples and sequences collected with funding from USAID-PREDICT, particularly in Yunnan.
- **Note:** blocking access to these samples and related sequences is a very effective obstruction of scientific efforts to investigate the origin of SARS-CoV-2. With this information scientists could build a better phylogeny tree for SARS-CoV-2 which could point to the likely time and location of the virus emergence. In particular:
  - The RdRp of bat betacoronavirus 7896 and almost identical RdRp sequences of seven other viruses have been published to GenBank. They are some of the most similar known viruses to SARS-CoV-2, but their Spikes have not been published.
  - Due to the presence of "7896" in the names of the RaTG13 amplicon sequences, we have reason to believe that RaTG13 may be a combination of BtCoV/4991 and viruses from the 7896 clade, rather than the full sequence of BtCoV/4991. Access to this database would allow us to verify this. It is not standard practice to name sequences purportedly from one virus after a closely related virus. As the 7896 clade appear to be further from SARS-CoV-2 than BtCoV/4991 is, using sequences from the 7896 clade in the RaTG13 consensus sequence rather than BtCoV/4991 sequences would be likely to make RaTG13 appear significantly less similar to SARS-CoV-2 than the full BtCoV/4991 sequence would have been.

Phylogenetic tree from a [video presentation by Shi Zhengli](#) showing the 'new' 7896 beta-BatCov clade: It seems that it is [based on more than the usual short RdRp sequence](#)



# Extended RdRp phylogenetic tree with 7896 class:

Built using [data](#) from the [Latinne et al paper](#) (anonymous contribution)



## Sampling information on the 7896 clade from Yunnan:

Viruses	Accessi	Genera	Species	Cite	Month	Group	sam	cave prov	collecti	Modific	Prepar
Bat betacoronavirus isolate 7896	MN312671	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7896-2	Yunnan	May-15	01-Jun-20	13-Aug-19
Bat betacoronavirus isolate 7905	MN312632	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7905	Yunnan	May-15	01-Jun-20	
Bat betacoronavirus isolate 7907	MN312633	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7907	Yunnan	May-15	01-Jun-20	
Bat betacoronavirus isolate 7909	MN312634	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7909	Yunnan	May-15	01-Jun-20	
Bat betacoronavirus isolate 7921	MN312635	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7921	Yunnan	May-15	01-Jun-20	
Bat betacoronavirus isolate 7924	MN312636	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7924	Yunnan	May-15	01-Jun-20	
Bat betacoronavirus isolate 7931	MN312672	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7931-2	Yunnan	May-15	01-Jun-20	
Bat betacoronavirus isolate 7952	MN312673	Betacoronavirus	SARS-like	Latinne et al. (2020)	Aug-20	WIV+EcoH	7952-2	Yunnan	May-15	01-Jun-20	

Blasting of [Ra7952](#) RdRp showing similarity between the 8 members of the 7896 clade (MN3123xx):

Sequences producing significant alignments									
Download <span>▼</span> <span>New</span> Select columns <span>▼</span> Show <span>10</span> <span>▼</span> <span>?</span>									
<input checked="" type="checkbox"/> select all	10 sequences selected								
	<a href="#">GenBank</a>	<a href="#">Graphics</a>	<a href="#">Distance tree of results</a>	<span>New</span> <a href="#">MSA Viewer</a>					
	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7952 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	743	743	100%	0.0	100.00%	402	<a href="#">MN312673.1</a>
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7931 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	737	737	100%	0.0	99.75%	402	<a href="#">MN312672.1</a>
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7896 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	730	730	99%	0.0	99.50%	591	<a href="#">MN312671.1</a>
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7924 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	730	730	99%	0.0	99.50%	589	<a href="#">MN312636.1</a>
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7909 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	730	730	99%	0.0	99.50%	592	<a href="#">MN312634.1</a>
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7907 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	730	730	99%	0.0	99.50%	597	<a href="#">MN312633.1</a>
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7921 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	725	725	99%	0.0	99.50%	588	<a href="#">MN312635.1</a>
<input checked="" type="checkbox"/>	Bat betacoronavirus isolate 7905 RNA-dependent RNA polymerase (RdRp) gene, pa...	Bat betacoro...	682	682	93%	0.0	99.47%	566	<a href="#">MN312632.1</a>
<input checked="" type="checkbox"/>	Severe acute respiratory syndrome coronavirus 2 isolate 190017 genome assembly...	Severe acut...	603	603	99%	4e-168	93.77%	29903	<a href="#">LR862443.1</a>
<input checked="" type="checkbox"/>	Severe acute respiratory syndrome coronavirus 2 isolate 180051 genome assembly...	Severe acut...	603	603	99%	4e-168	93.77%	29903	<a href="#">LR861674.1</a>

### Finding: Questions as to the unavailability of batvirus.whio.ac.cn are not being properly answered

- In a [BBC interview](#) Pr. Shi Zhengli is on record explaining that the VIW database was offline due to security reasons:

She told the BBC that the WIV's website and the staff's work emails and personal emails had been attacked, and the database taken offline for security reasons.

However it is extremely difficult to imagine why a copy of the database (or at least of the non-password protected portion) could not be safely put back online, if necessary on a totally different network unconnected to the WIV.

- In the same interview, she adds that:

"All our research results are published in English journals in the form of papers," she said. "Virus sequences are saved in the [US-run] GenBank database too. It's completely transparent. We have nothing to hide."

Her statement is demonstrably false. As mentioned before Peter Daszak, president of EcoHeath Alliance which was involved in the EPT samplings with the WIV, said just [as much in an interview](#):

"[...] we have an ongoing collaboration, [we have data that we've gathered over 15 years of working in China — 5 years under a previous grant from the NIH — which haven't been published yet](#). So we need to carry on with that work."

And as pointed previously: Batvirus.whio.ac.cn holds a password protected section for data that was not to be freely shared. That private data includes:

- (1) Sampling data of viruses not yet sequenced
  - (2) Sequences of viruses that have not yet been the object of publication.
  - (3) Sampling data from sites that the WIV prefers not to make public.
- Two separate direct enquiries to the WIV by members of the DRASTIC team (details are confidential) were met with inconsistent answers:
    - One answer gave cyber security issues as the reason for a temporary unavailability.
    - Another answer simply mentioned that the database was unavailable because it was being updated.
  - Questions to [CSDData](#) which hosted the description of the database and a live link to hit, as to why the link was not working, have not been answered.

**Finding: The suspension of the availability of batvirus.whiov.ac.cn is contrary to Chinese law**

---

The legal requirements and conditions for the publications of Scientific data are set in the 'Notice of the General Office of the State Council on Printing and Distributing Scientific Data Management Measures' ([Guobanfa \(2018\) No. 17](#)).

Of particle relevance are:

**Article 19**

The scientific data formed by government budget funding shall be organized in accordance with the principle of opening as the norm and not being the exception.

The competent department shall organize the compilation of the scientific data resource catalogue. The society and relevant departments are open to sharing, and the military-civilian sharing channels for scientific data are unblocked. Except for special provisions in national laws and regulations.

**Article 20**

The legal entity shall classify and classify scientific data, clarify the confidentiality level and confidentiality period, opening conditions, open objects and review procedures of scientific data, publish an open catalog of scientific data as required, and download, share or customize services through online And other ways to open to the community.

**Article 31**

For forgery of data, infringement of intellectual property rights, failure to submit data in accordance with regulations, etc., the competent department may order relevant units and responsible persons to order rectification, report criticism, punishment, etc., or give administrative penalties according to law.

Units and individuals violating relevant laws and regulations of the state shall be held accountable according to law.

Since the public part of batvirus.whiov.ac.cn was previously accessible, it is reasonable to conclude that it fell under the conditions for 'opening' of Article 19, and was not the subject of 'special provisions in national laws and regulations'. Once classified as non-confidential, it should be available as per Article 20, a breach of which is a violation of the law as per Article 31.

Hence the suspension of the availability of the public part of batvirus.whiov.ac.cn is a breach of the applicable Chinese law.





## 1.b Changes in the Description of the DB (30th Dec 2019):

The database description was separately hosted on the [CSData website](#), which is the main portal and index of Chinese scientific databases. The changes in the title and content of the database factsheets have been [highlighted in this document](#).

The database was also very likely used to answer the [national 2019 Scientific Survey project](#) started in July 2019 by the Chinese Ministry of Science & Technology. We note that the changes in the CSData description limit the scope of the database to Bats and Rodents, away from the previous larger 'Wild Animals' with also mentions of 'Insects'. It also removes a mention of cross-species infection. Overall this distances it from the 2019 Scientific Survey project.

**Example of changes** (refer to [comparison](#) for full changes):

VERSION 2 – 17 <sup>TH</sup> JULY 2019 (ZH2)	VERSION 4 – DECEMBER 30 <sup>TH</sup> 2019 (ZH4)
 <p>CN: 11-6035 / N ISSN: 2096-2223</p> <p>subscriptionPaper submissionhistory</p> <p>Scientific big data engineering <b>area I paper</b> (under review) • Version ZH2</p> <p><b>download</b></p> <p><b>Characteristic Database of Viruses Carried by Wild Animals</b></p> <p><b>Wildlife-borne Viral Pathogen Database</b></p> <p>Tang Yijie, Li Bei, Zhou Zijian, Zhu Yan, Zhao Kai, MA Li-li, Wuyue Wei, Shi Zhengli <sup>Corresponding</sup></p> <p>Tang Yijie, Li Bei, Zhou Zijian, Zhu Yan, Zhao Kai, Ma Lili, Wu Yuewei, Shi Zhengli <sup>Editor</sup></p> <p><a href="#">Visit the dataset &gt;&gt;</a></p> <p>DOI: 10.11922 / csdata.2019.0018.zh</p>	 <p>CN: 11-6035 / N ISSN: 2096-2223</p> <p>subscriptionPaper submissionhistory</p> <p>Scientific Big Data Engineering <b>Area II Paper</b> (Published) • Version ZH4 Vol 4 (4) 2019</p> <p><b>download</b></p> <p><b>Database of pathogens of bat and mouse origin</b></p> <p><b>Bat and rodent-borne viral pathogen database</b></p> <p>Tang Yijie, Li Bei, Zhou Zijian, Zhu Yan, Zhao Kai, MA Li-li, Wuyue Wei, Shi Zhengli <sup>Corresponding</sup></p> <p>Tang Yijie, Li Bei, Zhou Zijian, Zhu Yan, Zhao Kai, Ma Lili, Wu Yuewei, Shi Zhengli <sup>Editor</sup></p> <p><a href="#">Visit the dataset &gt;&gt;</a></p> <p>DOI: 10.11922 / csdata.2019.0018.zh PID: 21.86101.1 / csdata.2019.0018.zh</p>
<p>humans, predict hot spots of <b>emerging infectious diseases</b>, and provide theoretical basis and technical support for disease prevention and control.</p> <p><b>Keywords:</b> wild animal samples, viral pathogen data, emerging infectious diseases, cross-species infection</p> <p><b>Abstract &amp; Keywords</b></p> <p><b>Abstract:</b> Cross-species infection of viruses with wildlife origin is one of major reasons for emerging viral diseases in humans.</p> <p>Therefore, it is especially important to <b>construct database</b> for wildlife-borne viral pathogens based on collection and collation of wildlife samples and corresponding viral pathogen data.</p> <p>Through establishment of standard metadata description, this database includes up to 20,000 collated and processed data from <b>wildlife samples</b> collected from <b>many of countries</b> and several provinces and regions throughout China, and serves institutes and universities in China and other countries that conduct researches related to wildlife viral pathogens.</p> <p>The users can perform query based on wildlife species, sample type, species distribution, etc., according to their requirement.</p>	<p><b>infectious diseases</b>, and provide theoretical basis and technical support for disease prevention and control.</p> <p><b>Keywords:</b> bat, rat, virus, database</p> <p><b>Abstract &amp; Keywords</b></p> <p><b>Abstract:</b> Cross-species infection of viruses with wildlife origin is <b>one of the major origins for emerging viral diseases in mankind</b>.</p> <p>Therefore, it is especially important to <b>build a database</b> for wildlife-borne viral pathogens based on collection and collation of wildlife samples and corresponding viral pathogen data.</p> <p>Through establishment of standard metadata description, this database includes up to 20,000 collated and processed <b>data from bat and rodent samples</b> collected from <b>many countries</b> and several provinces and regions throughout China, and serves institutes and universities in China and other countries that conduct researches related to wildlife viral pathogens.</p> <p>The users can perform query based on bat and rodent species, sample type, species distribution, etc., according to their need.</p>

To understand how these changes may relate to the 2019 Scientific Survey project, here is the relevant section of that project description where the precise words that were removed from the database factsheet on the 30<sup>th</sup> Dec 2019 are highlighted in red.

## 15. Scientific investigation of virus carriers of main natural epidemic sources in China

Work content: focus on the systematic background screening of virus pathogens carried by **major wild animals** (bats, mice, birds, etc.) and **insect vectors** (ticks, mosquitoes) in China; investigate the range of virus carried by major animal hosts and vectors, and determine the natural geographical extent in the nation. The distribution of sexually potent pathogens; the use of biosafety platforms to isolate and identify new viruses; to conduct pathogenic basic characteristics research on important pathogenic new viruses; establish a national strategic virus pathogen resource library and shared database, and carry out resource and information sharing services.

Evaluation indicators: (1) Submit a survey report on the lineage, genetic characteristics and geographic distribution of bats, mice, birds, ticks, mosquitoes and other virulent virus pathogens in key areas in China; (2) Obtain more than 100 new viruses/strains Genome; (3) Complete the isolation and identification of 50 important viral pathogens and standardized collection and preservation; (4) **Complete the pathogenic characteristics analysis of 10 important new viruses/strains, which for at least 5 important pathogenic new viruses should be based on a biosafety risk assessment at the cell level and small animal level;** (5) Establish a standardized virus pathogen resource library and shared database.

## 1.c Importance of the research objectives of the 2019 Scientific Survey:

### 1.c.1 Likely work on undisclosed viruses:

As already mentioned, the 15.4 objectives above concern new viruses, so typically viruses on which no research has been published yet, and therefore most likely held in the password-protected section of the DB.

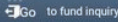
As also mentioned, good candidates could be found within the 8 beta BatCoV viruses of the 7896 clade which are close relatives of RaTG13 - for some reason still not fully sequenced and not the object of any publication by the WIV.

### 1.c.2 Biosafety risk assessment at the cell level:

This typically means introducing a virus - be it natural viruses, or chimeric [formed by combination of genetic material of some viruses] - into human and other animal cell lines to see if they replicate or attach to human ACE2 in vitro.

### 1.c.3 biosafety risk assessment at the small animal level:

This typically means work on humanized mice expressing the human ACE2 protein. Such pathogenicity studies have been going on for years at the WIV by Pr. Shi Zhengli and others at the WIV such as Ben Hu. See [example](#) (English [translation](#)).

<b>MedSci 梅斯</b> News    guide    tool    service    Coll    Please enter a keyword    login    registered    advisory	
<b>Pathogenicity of two new bat SARS-related coronaviruses to transgenic mice expressing human ACE2</b> Person in charge: Hu Ben Supporting unit: Wuhan Institute of Virology, Chinese Academy of Sciences Year of approval: 2018	
	
Project Description	
project name	Pathogenicity of two new bat SARS-related coronaviruses to transgenic mice expressing human ACE2
Project approval number	31800142
Subject classification	C010802 Life Sciences _ Microbiology _ virology _ animal virology
Type of funding	Youth Science Fund Project
principal	Hu Ben
Supporting unit	Wuhan Institute of Virology, Chinese Academy of Sciences
Year of approval	2018
Start and end time	201901-202112
Approved amount	250 thousand yuan
Summary	No data

Some pathogenicity research may also involve chimera viruses where various genetic variations of a natural backbone (simulating possible recombinations in nature) are studied. Such research was pioneered by Dr Ralph Baric (who is [on record](#) as saying that only the WIV archives can clarify the origins of Covid-19) and was commonly carried out at the WIV.

A good summary of these research directions is available [here](#):



Probably the largest reported number of novel chimeric viruses created was described in a 2017 paper from the Shi group at WIV, in which the authors reported creating eight chimeric viruses using WIV1 as a backbone and transplanting into it various RBDs from bat SARS-like viruses. These viruses were collected over a span of 5 years from the same cave near Kunming, Yunnan Province, where the Shi group originally found Rs3367 and RsSHC014. Only two of the eight live chimeric viruses were successfully rescued, and those two strains were found to possess the ability to bind to the human ACE2 receptor, as confirmed by experiments in hACE2-expressing HeLa cells and RT-PCR quantification of viral RNA.


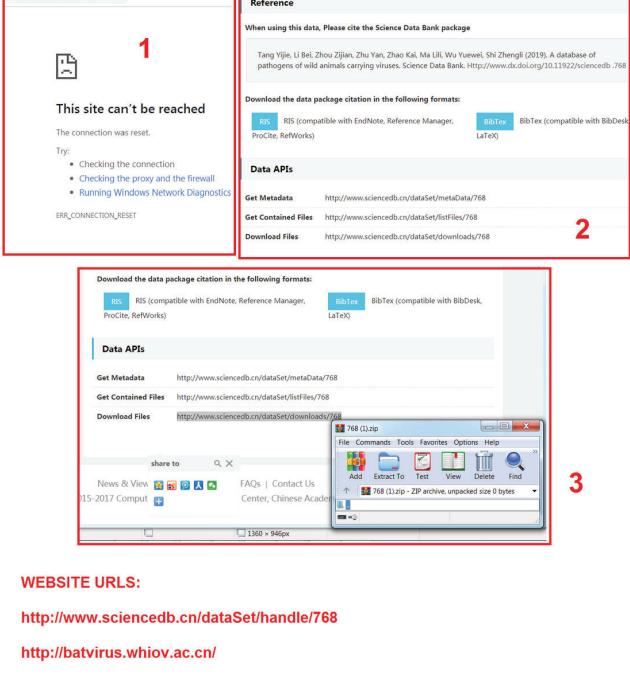
## 1.d Database Events

### 1.d.1 Timeline:

- **Apr 2019:** database version 1 [is released](#) and is accessed on a near daily basis
- **Jun 2019:** database version 2 is released.
  - It includes a password protected section for yet unpublished virus sequences.
  - [6.4GB](#) was downloaded from the 61.5MB SQL database from [within the China Science and Technology Network](#), (CSTN), nearly all [from Beijing](#), equivalent to 100 DB downloads.
- **16<sup>th</sup>- 22<sup>nd</sup> Aug 19:** odd [nil usage](#)
- **12<sup>th</sup> Sep 2019:** database seems [to be taken offline](#) (for non-WIV access), [see also this link](#). No reason for the gap and then the removal.
  - There is no access anymore either via <http://batvirus.whiov.ac.cn> or <http://www.sciencedb.cn/dataSet/handle/768>
  - Database has not been put back online (for external access) since then.
  - In [September](#) it was accessed almost entirely by the WIV and non-CSTN users. This indicates that the WIV were accessing data on viral pathogens from the program that identified RaTG13 well into September 2019. A full breakdown of the access records for September is available [in this table](#), based on data gathered from the [Chinese Scientific Database monitoring portal](#).
- **Since 12<sup>th</sup> Sep 2019:** records show the database being [online very intermittently](#) until at least early 2020, presumably without a DNS address and possibly only accessible from the WIV.
- **30<sup>th</sup> Dec 2019:** database title and description are modified - the day the [coronavirus breakout became public](#).

## 1.d.2 Zoom: Access Interruption

Both <http://batvirus.whio.ac.cn> and <http://www.sciencedb.cn/dataSet/handle/768> are now irresponsive.

Before 12 <sup>th</sup> Sep 19 (snapshot from June 19)	After 12 <sup>th</sup> Sep 19
	

batvirus.whio.ac.cn points to [159.226.126.81](http://159.226.126.81) which happens to be the same machine as for [www.viruses.nsd.cn](http://www.viruses.nsd.cn) ([159.226.126.6](http://159.226.126.6)).

```
PING batvirus.whio.ac.cn (159.226.126.81): 56 data bytes
Request timeout for icmp_seq 0
Request timeout for icmp_seq 1
Request timeout for icmp_seq 2
Request timeout for icmp_seq 3
Request timeout for icmp_seq 4
Request timeout for icmp_seq 5
Request timeout for icmp_seq 6
Request timeout for icmp_seq 7
^C
--- batvirus.whio.ac.cn ping statistics ---
9 packets transmitted, 0 packets received, 100.0% packet loss

PING www.viruses.nsd.cn (159.226.126.6): 56 data bytes
Request timeout for icmp_seq 0
Request timeout for icmp_seq 1
Request timeout for icmp_seq 2
Request timeout for icmp_seq 3
Request timeout for icmp_seq 4
Request timeout for icmp_seq 5
Request timeout for icmp_seq 6
^C
--- www.viruses.nsd.cn ping statistics ---
8 packets transmitted, 0 packets received, 100.0% packet loss
```

## 2. WIV: Sample resource management portal and DBs - [www.virus.org.cn](http://www.virus.org.cn)

### 2.a Summary:

Specifically:

- [www.virus.org.cn](http://www.virus.org.cn) is a portal of 15 databases, all of which are now offline as well as the portal
- Most likely these DBs are either derived from or form part of [batvirus.whiov.ac.cn](http://batvirus.whiov.ac.cn).

Then as already established for [batvirus.whiov.ac.cn](http://batvirus.whiov.ac.cn):

- These WIV DBs likely hold essential information on SARS-CoV-2 origins
- Partial funding via USAID-PREDICT of sample collections that inform some of these DBs is at odds with their unavailability
- Partial funding of WIV research by the NIH that may inform some of these DBs is at odds with their unavailability
- Financial support and membership of EVAg is at odds with the unavailability of [batvirus.whiov.ac.cn](http://batvirus.whiov.ac.cn) and the lack of transparency of the WIV

### 2.b Description:

The [www.virus.org.cn](http://www.virus.org.cn) was a product of the National Virus Resource Center (NVRC, 国家病毒资源库), which is affiliated to the WIV. It was essentially a WIV web portal which gave access to several key databases.

<http://www.viruses.nsd.cn> (mentioned in <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6178075/>) and [www.virus.org.cn](http://www.virus.org.cn) seem to have been pointing to the same portal.

It is quite likely that these DBs are either derived from or form part of [batvirus.whiov.ac.cn](http://batvirus.whiov.ac.cn).

As per its former 'About' page:

#### **Microbial Strain Collection (China Common Virus Collection):**

1. The Center for the Collection of Microorganisms and Viruses (VCCC) and the "China Common Virus Collection (CCGVCC)", "The Center for Common Microbiology (Virus) of the Chinese Microbial Culture Collection Management Committee", and the "Chinese Academy of Sciences Type Culture Collection Committee Virus Library"
2. Founded in 1979, it is affiliated to the Wuhan Institute of Virology, Chinese Academy of Sciences. It was registered with the International Culture Collection Committee (WFCC) in 1989 and was included in the 6 national collections of pathogenic microorganisms (viruses) for human infection by the National Health and Family Planning Commission in 2013. One of the institutions: the Microbial Strain Collection Center of Wuhan Institute of Virology, Chinese Academy of Sciences.
3. It is a comprehensive research center integrating virus resource collection and preservation, virus biotechnology research and development, systematic virology and bioinformatics research.
4. With its significant advantages in the variety, large number, and high level of biosafety of preserved virus resources, it has the largest virus depository in Asia, with approximately 1,200 strains of various virus isolates. The live viruses stored in the library cover human medical viruses, zoonotic viruses, animal viruses, insect viruses, plant viruses, bacteriophages, environmental microorganisms, virus-sensitive cell banks, and virus genetic resource banks.
5. It has created the country's only "Chinese Virus Specimen Museum" with modern display methods, integrating discipline, characteristics and popular science. It is the first batch of "National Youth Science and Technology Activity Demonstration Base".
6. **Supported by the European Union's H2020 plan**, it is one of the core cooperation units of the European Virus Resource Bank-Going Global (EVAg).

Note that the 'China Common Virus Collection' and the related database were part of EVAg and supported by grants from the European Union under the Horizon 2020 plan (see [Finding](#) under '1. WIV: Bat and rodent borne viral pathogens - batvirus.whiov.ac.cn' for details of these grants).

The portal and all the databases are now offline. See [www.virus.org.cn/resource](http://www.virus.org.cn/resource) for a capture of the site before it went offline. As per its description in the top right of that above page:

The live viruses preserved in the library include human medical viruses, zoonotic viruses, livestock and poultry viruses, insect viruses, plant viruses and bacteriophages. A virus-sensitive cell bank and a virus genetic resource bank have been established. A digital management system has been implemented, **13 virus databases and information sharing platforms have been established** and web services have been established. Over the years, the center has provided domestic and foreign research institutions, colleges and universities with related virus resources services such as teaching and research materials, and has maintained close contact and cooperation and exchanges with relevant domestic and foreign institutions.

(Incidentally that page listed bat-coronaviruses as BSL-2 level - possibly for cell culture only, not clear).

A 2016 entry in a Chinese catalogue gives some details of the number and types of viruses held at the China Center for General Viruses Culture Collection (CCGVCC), part of the WIV.

See [http://www.wfcc.info/ccinfo/collection/by\\_id/613](http://www.wfcc.info/ccinfo/collection/by_id/613):

Acronym	CCGVCC	
Full Name	China Center for General Viruses Culture Collection	
Institution	Wuhan Institute of Virology, Chinese Academy of Sciences	
Number of Strains until now		
Type	Holding strains new species	(total number)/(number from your country)
Cell lines: animal	110	110
Viruses: animal	922	922
Viruses: plants	160	160
Viruses: Bacteria	84	0
Viruses: Insect	135	0
Date of Entry	1981-10-14	
Date of Reply	2016-11-04	

There is also a good example of mention of these databases in a [paper by Yuan Zhiming](#) ('Investigation of Viral Pathogen Profiles in Some Natural Hosts and Vectors in China', Virologica Sinica, Mar 2018).

## 2.b Funding:

Here is another mention of the portal with a [brief description](#)..



The reference to the 11<sup>th</sup> plan (2006-2010) likely refers to the inception of the project as the databases metadata pages ([example](#)) show that the first DBs were set up in 2011, so about a year after initial funding.

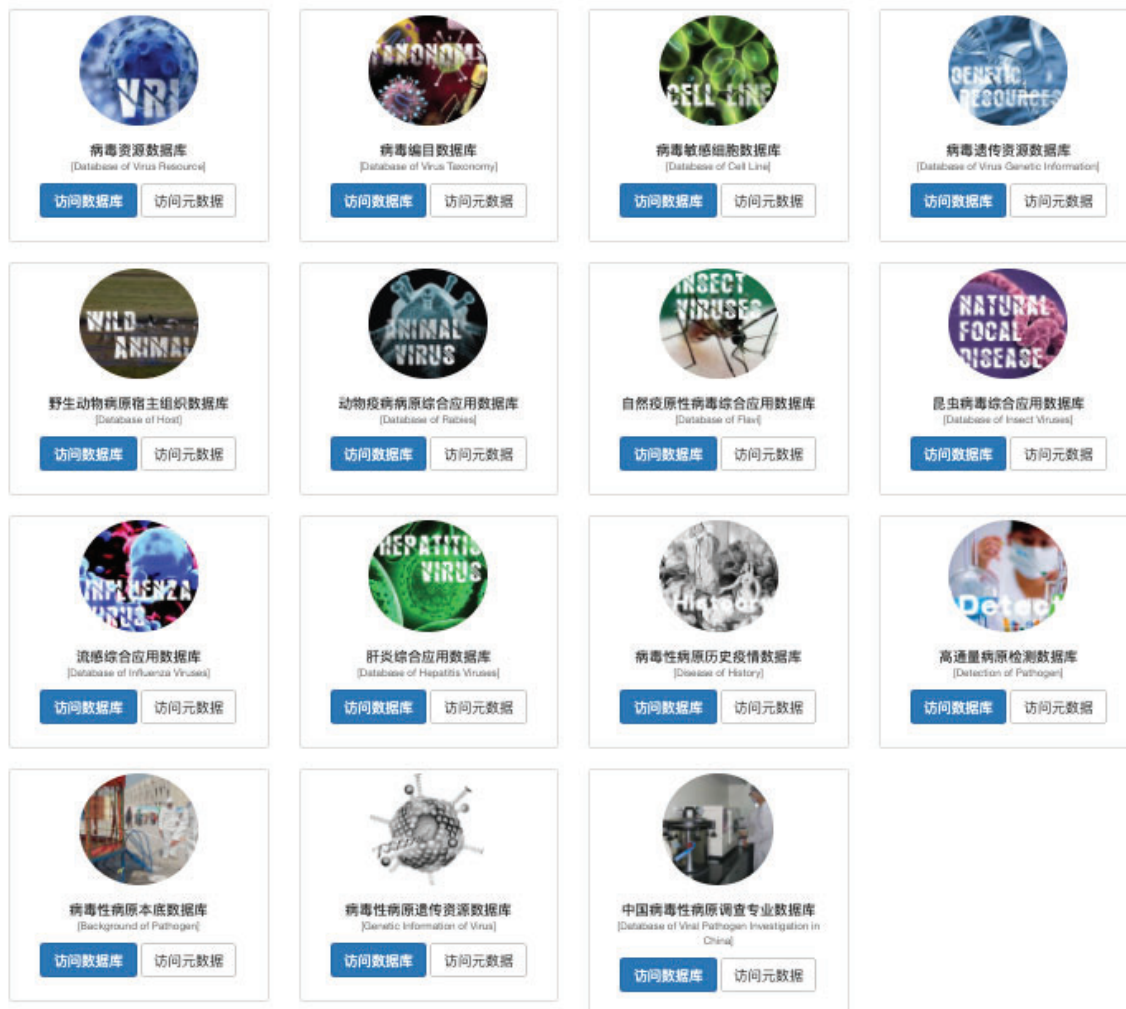
After its inception as part of the 11<sup>th</sup> Five-Year Informatization Construction Project, these DBs and the portal have been further funded via specific projects under the five-year plans that followed.

The latest of such projects is the 'Special Scientific Big Data Project of the Chinese Academy of Sciences' "[13<sup>th</sup> Five-Year Special Scientific Big Data Project](#)" ([translation](#)), part of the 13<sup>th</sup> Five Year Plan (2016-2020). That Special Scientific Big Data project was approved in [Feb 2016](#) and kicked off in late 2017 and called for work completion by the 31<sup>st</sup> Dec 2020. It was reviewed in [July 2019](#).



## 2.c Underlying DBs:

The portals at [www.virus.org.cn](http://www.virus.org.cn) and [www.viruses.nsd.cn](http://www.viruses.nsd.cn) gave access to 15 databases in all.





Document prepared by DRASTIC - to be used conditional on clear attribution to DRASTIC

### List of underlying DBs:

Databases under <a href="http://www.virus.org.cn">www.virus.org.cn</a> / <a href="http://www.viruses.nsd.cn">www.viruses.nsd.cn</a> portal (itself offline)					
#	name	description	URI /URL	Status	Issuing agency
1	Database of Virus Resource / 病毒资源数据库 <a href="#">metadata</a>	[A catalogue of virus databases] The virus resource database integrates databases with independent intellectual property rights, including virus preservation database, virus species database, type specimen database, virus species database, and establishes a large-scale "Chinese virus resource database". Its virus resources cover various virus databases, including Human medical virus database, animal virus database, zoonotic virus database, wild animal virus database, natural foci virus database, emerging infectious disease pathogen database, insect virus database, plant virus database, bacterial virus database, virus genetic resource database (Preserve important viral gene fragment cloning plasmid library, gene sequence library, important genetic material)	cn.csdb.viruses.w ww.vri <a href="http://www.viruses.nsd.cn/vri.jsp">http://www.viruses.nsd.cn/vri.jsp</a>	<b>offline</b>	Wuhan Institute of Virology, Chinese Academy of Sciences
2	Database of Virus Taxonomy / 病毒编目数据库 <a href="#">metadata</a>	The virus taxonomy database mainly contains basic information of more than 7,800 strains of viruses, including the English name, Chinese name, classification, host, collection time, location, source, quantity, original literature, biological safety level and other data information of the virus, as well as the physical and chemical information of the virus. Characteristics, genome information, virus classification information, to maximize the background information of the species; the establishment of the database provides a convenient and friendly search interface, which can be based on the classification level of the virus, such as the Chinese and English names of the family name and the species name Or part of the fields to query virus information.	cn.csdb.viruses.w ww.bdbm <a href="http://www.viruses.nsd.cn/bdbm.jsp">http://www.viruses.nsd.cn/bdbm.jsp</a> .bdbm /	<b>offline</b>	
3	Database of Cell Line / 病毒敏感细胞数据库 <a href="#">metadata</a>	The 'virus-sensitive cell line' database mainly collects and preserves the corresponding virus-sensitive cell lines (lines) resources of humans and animals in China; researches and develops new cell culture technologies, and researches and develops new technologies for the preservation and quality control of cell lines (lines); facing the whole country, Provide standardized cell lines (lines) and related services for research in the field of virology and biotechnology in my country.	cn.csdb.viruses.w ww.cell <a href="http://www.viruses.nsd.cn/bdmgxb.js">http://www.viruses.nsd.cn/bdmgxb.js</a> p	<b>offline</b>	
4	Database of Virus Genetic Information /	The virus genetic resource database (genetic resources-clone library) collects basic information about the cloning of important functional genes of various pathogens, including the biological background of the cloned genes, gene	cn.csdb.viruses.w ww.nuc	<b>offline</b>	
					Contact name: Fan Zhaojun Chinese Academy of Sciences Wuhan Institute of Virology No. 44, Xiaohongshan Middle District, Wuchang, Wuhan Phone 027-87199112 EMAIL roy@wh.iov.cn

Document prepared by DRASTIC - to be used conditional on clear attribution to DRASTIC

	病毒遗传资源数据库 <a href="#">metadata</a>	description, coding protein, function, references, etc.; and establishes and Reserve the nucleic acid detection methods for these pathogens.	<a href="http://www.viruses.nsd.cn/byczyklk.jsp">http://www.viruses.nsd.cn/byczyklk.jsp</a>		
5	Pathogenic Wild Animal Host Database / 野生动物病原宿主组织数据库	The collection of pathogenic wild animal host database includes different pathogen transmission vectors (rodents, birds, bats, mosquitoes, ticks), different sample types (throat swabs, anal swabs, serum, pathological tissues) storage temperature, storage methods, Scientific data such as physical status, sharing method, pathological background, etc	cn.csdb.viruses.www.host <a href="http://www.viruses.csdb.cn/vsdw.jsp">http://www.viruses.csdb.cn/vsdw.jsp</a>	<b>offline</b>	
6	Database of Animal Disease Pathogens / 动物疫病病原综合应用数据库	The comprehensive application database of animal disease pathogens integrates animal disease pathogen data and related virus strain information, and provides independent research epidemiological survey data, including: animal disease pathogen isolates information, biological characteristics, antigenic phenotype, pathogenicity, Genomics, genetic stability and molecular evolution characteristics, epidemiological analysis and geographic display systems, etc.; can intuitively realize the graphical display of geographic information of biological information systems such as molecular epidemiological data and epidemic dynamics; can integrate common analysis software, such as Multi-sequence comparison, sequence mutation analysis, evolution analysis, etc., to realize the safe storage and sharing of molecular data; it is also possible to conduct inference and analysis of the pathogen gene library based on the existing virus gene sequence and the evolution model based on the evolutionary model, thereby providing virology Related research provides a complete information service platform with convenient operation interface, safe data sharing, and retrieval content covering basic and professional databases.	cn.csdb.viruses.www.rabiesdb <a href="http://appdb.viruses.nsd.cn/rabiesdb">http://appdb.viruses.nsd.cn/rabiesdb</a>	<b>offline</b>	
7	Database of natural foci of viruses / 自然疫源性病毒综合应用数据库	The comprehensive application database of natural foci of viruses integrates natural foci virus data and related strain information, and provides independent research epidemiological survey data, including: natural foci virus isolate information, biological characteristics, antigen table Type, pathogenicity, genomics, genetic stability and molecular evolution characteristics, epidemiological analysis and geographic display system, etc.; can intuitively realize the graphical display of geographic information in biological information systems such as molecular epidemiological data and epidemic situation; Integrate common analysis software, such as multiple sequence comparison, sequence mutation analysis, evolution analysis, etc., to realize the safe storage and sharing of molecular data; it can also be carried out based on the evolutionary model based on the existing natural epidemic-borne virus gene sequence Pathogen gene bank inference and analysis; provide a convenient operation interface for virology related research, data sharing safety, and a	cn.csdb.viruses.www.flavdb <a href="http://appdb.viruses.nsd.cn/flavdb/">http://appdb.viruses.nsd.cn/flavdb/</a>	<b>offline</b>	

*Document prepared by DRASTIC - to be used conditional on clear attribution to DRASTIC*

		complete information service platform covering basic databases and professional databases			
8	Database of Insect Viruses / 昆虫病毒综合应用数据库	The insect virus comprehensive application database integrates insect virus data and related virus strain information, and provides independent research epidemiological survey data, including: natural foci virus isolate information, biological characteristics, antigenic phenotype, pathogenicity, Genomics, genetic stability and molecular evolution characteristics, epidemiological analysis and geographic display systems, etc.; can intuitively realize the graphical display of geographic information of biological information systems such as molecular epidemiological data and epidemic dynamics; can integrate common analysis software, such as Multi-sequence comparison, sequence mutation analysis, evolutionary analysis, etc., to realize the safe storage and sharing of molecular data; it can also perform pathogenic gene library inference and analysis based on the existing insect virus gene sequence based on the evolution model; for virology Related research provides a complete information service platform with convenient operation interface, safe data sharing, and retrieval content covering basic and professional databases	cn.csdb.viruses. www.bac <a href="http://appdb.virus.es.nsidc.cn/baculodb/">http://appdb.virus.es.nsidc.cn/baculodb/</a>	<b>offline</b>	
9	Database of Influenza Viruses / 流感综合应用数据库	The Influenza Comprehensive Application Database integrates influenza virus data and related strain information, and provides independent research epidemiological survey data, including: influenza virus isolate information, biological characteristics, antigenic phenotype, pathogenicity, genomics, genetics Stability and molecular evolution characteristics, epidemiological analysis and geographic display system, etc.; it can dynamically realize the graphical display of geographic information in the biological information system of influenza virus, including: pathogen, circulation information, time and space information, transmission route, infection mode, Susceptible host, range, etc.; it can integrate common analysis software, such as multiple sequence alignment, sequence mutation analysis, evolution analysis, etc., to realize the safe storage and sharing of molecular data; it can also be based on the existing influenza virus gene sequence. Based on the evolution model, the pathogen gene bank is inferred and analyzed; it provides a convenient operation interface for virology-related research, data sharing is safe, and a complete information service platform covering basic databases and professional databases; for influenza virus epidemic early warning and prevention	cn.csdb.viruses. www.aiv <a href="http://appdb.virus.es.nsidc.cn/influb/">http://appdb.virus.es.nsidc.cn/influb/</a>	<b>offline</b>	

Document prepared by DRASTIC - to be used conditional on clear attribution to DRASTIC

		Control provides a multi-level information exchange platform.			
10	Database of Hepatitis Viruses / 肝炎综合应用数据库	The Hepatitis Comprehensive Database integrates various hepatitis virus data and related strain information, and provides epidemiological survey data for independent research.	<a href="http://www.virusen.s.nsdcc.cn/hepatitis.jsp">http://www.virusen.s.nsdcc.cn/hepatitis.jsp</a>	<b>offline</b>	
11	Disease of History / 病毒性病原历史疫情数据库	Contains a series of historical epidemic data including the outbreak of key pathogens. The specific data should include the time of the outbreak, the end time, the spread of the epidemic, the number of infections, and the number of deaths (rate), and restore as much as possible the habitat data of the foci at the time of the outbreak. At the same time, the secondary data processing is carried out, and the historical epidemic database is established according to different regions, different hosts, different pathogens, and habitats.	<a href="http://www.virusen.s.nsdcc.cn/history.jsp">http://www.virusen.s.nsdcc.cn/history.jsp</a>	<b>offline</b>	
12	Detection of Pathogen / 高通量病原检测数据库	Through bioinformatics analysis, establish the experimental technology of Pan-PCR and multiple primer combined PCR for viruses carried by different species (bats, birds, mice, ticks, and mosquitoes); and establish a laboratory for bats, birds, mice, ticks, and mosquitoes. Nucleic acid detection methods on a flow-based liquid chip that carry important viruses; develop new rapid detection methods for nucleic acids, antigens, and antibodies that carry viruses in natural hosts and vector insects; use the above methods to perform high-throughput virus detection on samples collected in different regions. Contains the result data of metagenomic sequencing and analysis of related samples to construct a high-throughput pathogen detection database.	<a href="http://www.virusen.s.nsdcc.cn/detect.jsp">http://www.virusen.s.nsdcc.cn/detect.jsp</a>	<b>offline</b>	
13	Background of Pathogen / 病毒性病原本底数据库	Enter the detection and analysis data of the collected samples carrying viral pathogens to form the investigation report of different species carrying viral pathogens in the representative natural focus areas (Xinjiang, Qinghai, <b>Hubei</b> , <b>Yunnan</b> ), and obtain the typical natural focus of viral diseases in my country. The new virus background of China, the construction of the original database of Chinese viral diseases, provides important data for major basic research such as understanding the interaction between host animals, vector insects, and viruses, and revealing the outbreak mechanism of new and emerging infectious diseases.	<a href="http://www.virusen.s.nsdcc.cn/pathogen.jsp">http://www.virusen.s.nsdcc.cn/pathogen.jsp</a>	<b>offline</b>	

*Document prepared by DRASTIC - to be used conditional on clear attribution to DRASTIC*

14	Genetic Information of Virus / 病毒性病原遗传资源数据库	Contains the isolation and identification data of viral pathogens in different animal species and the information of new pathogen isolates, and establish a database of viral genetic resources from bats, mosquitoes, ticks, rodents, and birds in my country, providing indispensability for major scientific research such as vaccines and drugs Indispensable source resources.	<a href="http://www.virus.s.nsd.cn/pathogen.jsp">http://www.virus.s.nsd.cn/pathogen.jsp</a>	<b>offline</b>	
15	Database of Viral Pathogen Investigation in China / 中国病毒性病原调查专业数据库	This project integrates the original database on the basis of China's virus resource basic database. Aiming at the "Investigation of Viral Pathogens of my country's Important Natural Hosts and Vector Insects", the stored pathogen survey will collect new virus background data and genetic resources, which is a new Provide data for major basic research such as the outbreak mechanism of sudden infectious diseases, and will also provide indispensable source data resources and standards for major scientific research such as early diagnosis of emerging infectious diseases, vaccines and drugs.	<a href="http://www.virus.s.nsd.cn/chinavpi">http://www.virus.s.nsd.cn/chinavpi</a>	<b>offline</b>	

## 2.d Zoom: Access Interruption:

The portals and all the databases are inaccessible. The machine is there but not serving any web-page at any of the relevant addresses.

```
PING www.virus.org.cn (39.99.133.234): 56 data bytes
64 bytes from 39.99.133.234: icmp_seq=0 ttl=36 time=330.017 ms
64 bytes from 39.99.133.234: icmp_seq=1 ttl=36 time=328.971 ms
64 bytes from 39.99.133.234: icmp_seq=2 ttl=36 time=329.622 ms
64 bytes from 39.99.133.234: icmp_seq=3 ttl=36 time=329.284 ms
64 bytes from 39.99.133.234: icmp_seq=4 ttl=36 time=327.710 ms
```